University of Northern Iowa

UNI ScholarWorks

Sep 18th, 9:45 AM - 10:15 AM

Big Opportunities or Big Problems?: Participants' Views on Big Data

Helen Harton
University of Northern Iowa

Michael Mintz
University of Northern Iowa

See next page for additional authors

[Let us know how access to this document benefits you](#)

Recommended Citation

Author

Helen Harton, Michael Mintz, and Kristin Broussard

Big Opportunities or Big Problems?: Participants' Views on Big Data

Helen C. Harton, Michael Mintz, & Kristin A. Broussard

Department of Psychology

University of Northern Iowa

Abstract

"Big data" is a buzzword that refers to massive amounts of data being collected by websites and other entities that make traditional analyses difficult. There is excitement about what could be learned using big data, but there are also concerns about privacy and the rights of research participants, who often aren't even aware their data are being collected.

Introduction

More data about individuals are available in more forms than at any point in history. This shift in data availability has changed how researchers and laypersons think about data privacy and informed consent (Schadt, 2012). However, as researchers promote big data research, the average user has growing concern over how their privacy is being protected online.

Facebook recently published a study wherein they manipulated the newsfeeds of users without their explicit consent (Kramer, Guillory, & Hancock, 2014). Concern was raised over the ethics of the study, particularly how researchers could have better approached the consent process and reduced the perceived harm from privacy violations (Ross, 2014); however, going forward, researchers need to consider how new ethical concerns need to be considered in the digital age.

To more systematically ascertain how people feel about online companies using their data, we investigated how mTurk users (frequent study participants with online experience) and college students perceived research practices of companies that collect "big data" and how their own behaviors and beliefs about online security related to their concerns about how these companies were using their data.

Method

- 223 mTurk participants and 248 Introduction to Psychology students indicated how okay they were with each of several big data companies they used (i.e., Facebook, Twitter, email, dating websites, Amazon, local stores, GPS providers, and city data programs) using their data for internal versus published research of different types.
- Participants also indicated overall whether sites should:
  - ask permission before using their data;
  - provide information on what their data were used for after the study concludes;
  - compensate the people whose data were used; and
  - provide users with the option to either opt-in or opt-out of the study.
- Participants responded to questions about their concerns with online privacy and their data privacy behaviors using established scales.
- Participants rated how warmly they felt about specific companies and completed scales assessing trust in others and conspiratorial beliefs (all αs>.80).

Attitudes about Data Use by Type and Sample

Item	mTurk	College
Internal research*	2.87 (.83)	3.03 (.71)
Employee publication*	2.42 (1.01)	2.65 (.74)
Academic research without identifiers*	2.72 (.95)	2.90 (.73)
Academic research with identifiers*	1.94 (.99)	2.18 (.79)

Note. Means and (standard deviations) for how okay participants were with different data use contexts. Individual items were from rated from 1 (*strongly disagree*) to 5 (*strongly agree*). * All means different by rows at *p* < .05

How Okay Participants Were with Academic Use W/O IDs by Source

Item	mTurk	College
City data sensors*	3.28 (1.26)	3.61 (1.00)
Twitter basic info	3.17 (1.20)	3.37 (1.11)
Online shopping sites	3.09 (1.28)	2.95 (1.23)
GPS	3.05 (1.31)	3.15 (1.18)
Dating site basic info	2.98 (1.30)	3.48 (1.08)
Facebook basic info*	2.84 (1.35)	3.16 (1.20)
Twitter tweets	2.82 (1.30)	2.97 (1.21)
Local stores	2.82 (1.32)	2.97 (1.22)
Email providers basic info*	2.82 (1.31)	3.08 (1.18)
Dating site changing	2.56 (1.31)	2.90 (1.33)
Twitter changing	2.44 (1.18)	2.61 (1.09)
Facebook changing*	2.17 (1.20)	2.61 (1.02)
Facebook posts/photos	2.01 (1.27)	2.22 (1.20)
Email providers content	1.89 (1.25)	2.01 (1.20)

Note. * groups differ at *p* < .05. Changing = Changing what you see

Correlations of Individual Difference Variables and Attitudes

Measure	Internal	Employee	AR no ID	AR w/ID
Online safety practices	.01	-.01	.01	-.09
Concern about online security	-.09*	-.07	-.07	-.02
Warmth toward Big Data companies	.26**	.26**	.22**	.24**
Data privacy precautions	-.02	-.01	.01	-.01
Belief in conspiracies	-.15**	-.15**	-.16**	-.12*
General trust	.19**	.15**	.18**	.14**

Note. * *p* < .05, ** *p* < .01

Attitudes about Research by Sample

Item	mTurk	College
Ask permission for research*	4.47 (.86)	4.64 (.72)
Inform on purpose afterward*	4.24 (1.04)	4.35 (1.05)
Pay to use my data*	3.92 (1.08)	3.75 (1.10)
Allow opt-in/opt-out*	4.57 (.82)	4.66 (.67)

Note. Means and (standard deviations) for how okay participants were with different data use contexts. Individual items were from rated from 1 (*strongly disagree*) to 5 (*strongly agree*). * All means different by rows at *p* < .05

Results

- College students were generally more okay with their data being used for research, but neither group was very approving.
- People were generally more okay with data being used for internal research and academic research without identifiers than with employee publication or external research with identifiers.
- People were generally more okay with basic information being shared than with manipulations of their sites or content being used.
- Attitudes were not related to concerns about online safety.
- People who were more okay with their data being used for research had more positive attitudes toward big data companies and higher levels of general trust and tended to be less conspiracy-minded.
- Participants strongly agreed that they should have to give consent for their data to be used in research and that they should be allowed to opt-out of research.

Discussion

Although there is a push to use big data in research, many users are distrusting of companies using their data in this fashion. Our study shows that people generally do not want their data used in research, especially without their knowledge. Our next steps will investigate to what extent people are aware of how their data may be used and decide to use the sites anyway. Internet users, and especially those who may be older and more savvy, are distrusting of what they perceive to be violations of their privacy. Researchers should take steps to bring their ethical practices into line with how participants view their data rights.

References

Kramer, A. D. I., Guillory, J. E., & Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *PNAS: Proceedings of the National Academy of Sciences of the United States of America, 111,* 8788–8790. doi:10.1073/pnas.1320040111

Ross, M. W. (2014). Do research ethics need updating for the digital age? *Monitor on Psychology, 45,* 64. Retrieved from http://www.apa.org/monitor/2014/10/research-ethics.aspx

Schadt, E. E. (2012). The changing privacy landscape in the era of big data. *Molecular Systems Biology, 8,* 1-3. doi:10.1038/msb.2012.47